PL-TR-95-2017
Environmental Research Papers, No. 1165

SYSTEMATIC HUMIDITY ERRORS IN NUMERICAL WEATHER PREDICTION MODELS

Donald C. Norquist

13 February 1995

PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010

This technical report has been reviewed and is approved for publication.

DONALD A. CHISHOLM
Chief, Satellite Analysis and Weather
Prediction Branch

ROBERT A. McCLATCHEY
Director, Atmospheric Sciences Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	13 February 1995	Final Report

4. TITLE AND SUBTITLE

Systematic Humidity Errors in Numerical Weather Prediction Models

5. FUNDING NUMBERS

PE 61201F
PR IR3R
TA GT

WU 01

6. AUTHOR(S)

Norquist, Donald C.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Phillips Laboratory (GPAP)
29 Randolph Road
Hanscom Air Force Base, Massachusetts 01731-3010

8. PERFORMING ORGANIZATION REPORT NUMBER

PL-TR-95-2017
ERP, No. 1165

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In this study, we have examined the systematic humidity errors of both global meteorological analyses and a global numerical weather prediction (NWP) model (in various versions) initialized from them. Upon determining a definition for systematic error, we obtained the individual forecast bias and the ensemble bias over a sequence of forecast runs for each of several versions of the model. We found that model formulations designed to reduce systematic humidity forecast error with respect to one set of initial conditions did not produce the corresponding error reductions when initialized from a different set of meteorological analyses. Thus, both model formulation and initial conditions affect systematic humidity forecast error. Statistics gathered from several operational mesoscale NWP models suggest that each model has its own unique error characteristics.

14. SUBJECT TERMS

Numerical weather prediction — Systematic error
Humidity forecasting — Meteorological Analyses

15. NUMBER OF PAGES

40

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

Contents

Illustrations

Tables

Acknowledgement

We thank Dr. Kenneth Mitchell, of the National Meteorological Center, and Major James Millard, of Air Force Global Weather Central, for providing forecast error statistics. Data were provided by the European Centre for Medium Range Weather Forecasts and the United States Air Force Environmental Technical Applications Center. Ms. Audrey Campana accomplished the word processing. This research was funded by the Phillips Laboratory In-House Laboratory Independent Research program.

Systematic Humidity Errors in Numerical Weather Prediction Models

1. BACKGROUND

Numerous investigations of the performance of numerical weather prediction (NWP) models have been conducted in an attempt to find ways to reduce prediction errors. These errors generally fall into one of two categories. Departures in the forecast fields from verifying reference values are systematic errors if they are consistently of the same arithmetic sign in forecast after forecast, regardless of the weather scenario. Quantitatively, this means that the individual forecast mean error (IME) of the verification of a series of forecasts is always positive or always negative. The presence or absence of systematic error can be determined when an ensemble of forecasts is verified over a particular geographic region, altitude, and season. Random errors, measured in terms of mean absolute error [root-mean-square error (RMSE)], refer to an averaged magnitude of departure (mean-squared departure) from the reference. They are often the result of forecasts that misplace or produce the wrong intensity of weather patterns. Systematic errors are a contributor to the random error; for example the IME, or bias, is related to the RMSE as follows: RMSE =

Received for publication 31 January 1995

$\{ISDE^2 + IME^2\}^{1/2}$ for each forecast, where ISDE is the individual forecast standard deviation of the error. In efforts to improve a forecasting method, it is usually best to first try to reduce the systematic errors. The causes of the systematic errors in a complex forecast procedure like an NWP model are usually easier to determine and correct than the random errors. In addition, reducing the systematic error leads to a reduction in the RMSE.

Water vapor and cloudiness concentration forecasts have long been a top priority of the U.S. Air Force. It is probable that NWP models will ultimately play a major role in operational humidity and cloud amount forecasting in support of military support and planning. However, in comparison to mass and motion fields, assessment and improvement of NWP moisture forecasting has received relatively little attention in the operational and research communities. Perhaps the main reason for this inattention is the relative scarcity and inaccuracy of atmospheric water vapor and cloud water measurements worldwide. This makes moist atmospheric models difficult to reliably initialize and verify. Another possible reason for the lack of attention to moisture forecasting is the fact that water vapor or cloudiness forecasts are not required by the user in most operational centers. Instead, they have traditionally only been considered important for their effect on temperature and precipitation forecasts.

Though virtually all operational and research NWP models carry humidity (either absolute or relative) as a predictive variable, few studies document the verification of humidity forecasts. Those studies suggest that the models do generate a systematic error in the humidity forecasts. Heckley[1] found that the European Centre for Medium-Range Weather Forecasts (ECMWF) global model forecasts of tropical regions tended to produce an overly moist lower troposphere early in the forecast, then become drier in the late stages. Williamson[2] showed that the National Meteorological Center (NMC) global NWP model developed a deficit in specific humidity in 1000 and 400 mb Northern Hemisphere forecasts. Campana et al.,[3] reported that the NMC model generally underestimates cloud amounts when diagnosed from relative humidity (RH) forecasts, corroborating the dry bias seen by

Williamson.[2] Kuo et al.,[4] found that the Pennsylvania State University--National Center for Atmospheric Research (PSU-NCAR) Mesoscale Model version 4 (MM4) produced overly moist forecasts in the lower troposphere in forecast experiments over the U.S. central plains.

This report documents the continuation of a study of systematic humidity distribution errors in forecasts produced by NWP models. In an earlier paper[5] we reported on the evaluation of the systematic humidity errors in forecasts of the Phillips Laboratory Global Spectral Model (PL GSM)[6] and attempts to reduce them. The following paragraphs briefly summarize that paper's major findings. We began by defining systematic humidity forecast error. The forecast error in a particular region, altitude, and season is called *systematic* when the magnitude of the ensemble mean error of a series of forecasts of length n is greater than or equal to the standard deviation (SDE) of the individual forecast mean error (IME) about the ensemble mean error (ME). That is, the error of an n hour forecast of specific humidity (q) is systematic when $|ME_n| \geq SDE_n$, where

$$ME_n = \frac{1}{N_n} \sum_{j=1}^{N_n} IME_{j,n} \tag{1}$$

$$IME_{j,n} = \frac{1}{M_{j,n}} \sum_{i=1}^{M_{j,n}} (q^f_{i,j,n} - q^o_{i,j,n}) \tag{2}$$

$$SDE_n = \{\frac{1}{N_n} \sum_{j=1}^{N_n} (IME_{j,n} - ME_n)^2\}^{1/2}. \tag{3}$$

Here, q^f and q^o are the forecast and radiosonde observed (RAOB) q at each of M observation sites i in the region/pressure level category for the jth forecast of length n, and N_n is the number of forecasts of length n in the ensemble. Strictly speaking, the ensemble mean error is the average departure from the observed over all observations and forecasts, rather than the average of IME over all forecasts. We assumed that $M_{j,n}$ is essentially invariant from forecast to forecast, so that ensemble mean error is essentially as defined in Eq. (1). We sought to identify the

region/pressure level categories in the global humidity forecasts that were characterized by systematic errors according to this criterion.

We hypothesize that systematic humidity forecast errors were the result of incorrect initial humidity specification, deficiencies in the forecast model formulation, or both. We first evaluated the humidity distribution in the 1979 First GARP Global Experiment (FGGE) level IIIb global meteorological analysis data set. We found that for 0000 and 1200 UTC analyses, in both January and June, the specific humidities (computed from analyzed temperature and RH) were consistently too moist below 300 mb in all regions when compared with RAOB data.

We next attempted to remove the systematically overspecified specific humidity bias from the analysis data set by subtracting from each q analysis the difference between the monthly, zonal average of FGGE IIIb q, and that of a climatology[7] based on RAOBs for both months. This forced the resulting "corrected" analyses to have a zonal, time average equal to the RAOB-based climatology[7], while maintaining the patterns of the original longitudinal and time departures from the zonal, time average. A series of 18-layer, rhomboidal 30-wave PL GSM forecasts were initialized with both the original and corrected analyses. Verification results showed that, while 0 hour forecast q bias was much reduced in the corrected runs, the vertical profiles of q bias became very similar in shape and magnitude to the original bias profiles by two days into the forecast. Although the corrected humidity initial conditions led to improved temperature and precipitation forecasts, the PL GSM tended to produce similar vertical structures of systematic humidity error.

Our next step was to determine the region and pressure levels where systematic change of q was occurring in the model (PL GSM) forecasts. We used the corrected FGGE IIIb analyses to initialize the forecasts. The global domain was partitioned into three verification regions (90-30S, 30S-30N, and 30-90N) and the forecasts were verified against RAOBs on pressure levels. We found that PL GSM forecasts were characterized by systematic upper tropospheric moistening in all regions and by lower tropospheric drying in the tropics.

Next, we assessed the magnitude of each of the tendency terms in the model's

predictive equation for specific humidity. We looked for relationships between the sign and magnitude of the terms and the sign and magnitude of systematic errors we found in the humidity forecasts. We found that low-level tropical drying was linked to moist convection parameterization tendency, while the upper level moistening at all latitudes was attributable primarily to the horizontal and vertical advection terms.

We next sought alternative formulations for the PL-91 version of the PL GSM that would result in a reduction of systematic errors. These modifications were intended to address the dominant change terms in the model's specific humidity predictive equation. In experiment SFEVP, we adjusted the surface evaporation rate in the model by a constant factor to reduce it to the climatological precipitation rate[8] to force a hydrological balance. In experiment EMANU, we replaced the PL GSM's moist convective parameterization scheme[9] with a convective parameterization developed by Emanuel.[10] After conducting the experiments, we observed a large impact on water vapor change trends due to the alternative convective scheme. Because of this, we conducted the experiment PLKUO, in which we used the modified Kuo scheme of Norquist and Yang.[11] Finally, we modified the model's q predictive equation to use $Q = q/q_c$ as the prognostic variable, where q_c is the three-dimensional gridded monthly averaged (climatological) values of specific humidity. This experiment was termed QNORM, and was conducted to test the hypothesis that reduced horizontal and vertical gradients of the prognostic humidity variable lead to reduced transport of moisture down the gradient by horizontal and vertical advection.

The experimental five-day forecasts were initialized using the corrected FGGE IIIb analyses for 12 initial times spaced 2.5 days apart in June 1979. Forecasts of q were verified at 12-hour intervals against RAOB data on constant pressure levels in verification regions 90-30S, 30S-30N, and 30-90N. In each region/pressure level verification category at each 12-hour forecast interval, we computed the ensemble mean error (ME) of q, and the standard deviation (SDE) of the individual forecast mean error about the ensemble mean error as in Eqs. (1) and (3). Specific humidity errors were considered systematic for a given experiment in a region/pressure level category if $|ME| \geq SDE$ for at least seven of ten 12-hour interval forecast times. For

the experiment/region/pressure level categories where systematic error occurred, we computed the experimental mean error (EME) of specific humidity in percent bias (QP) as

$$QP = [\overline{(q^f - q^o)}/\overline{q^o}] \; X \; 100\% \tag{4}$$

where the overbar represents the average over all verifying RAOBS in the ensemble for a particular region and pressure level.

The results for the five June 1979 experiments conducted in the earlier study are reproduced here as Table 1. The primary observations made from the results are as follows. Use of the alternative convection parameterizations reduced systematic low-level tropical drying, but increased low- and middle-level moistening in the summer extratropics. The surface evaporation and alternative prognostic variable experiments revealed a reduced upper level moistening in the winter extratropics, and in one case reduced the low-level tropical drying (while at the same time increasing the upper-level tropical moistening). When we diagnosed the corresponding relative humidity error (using the same criteria for systematic error) as shown in Table 2, we found that systematic negative values were widespread in the region and pressure levels, due to a preponderance of warm bias in the model.

We concluded in our study that non-convective model modifications show promise of improving model humidity forecasts by reducing systematic humidity errors; therefore, there is incentive to further investigate their potential and refine their formulation, especially in improving the physical realism of the surface moisture flux parameterization. On the other hand, we concluded that moist convective parameterizations have shortcomings in their parameterization of sub-grid scale processes insofar as they impact the large-scale humidity distribution. More attention needs to be given to the convective and stratiform precipitation schemes' role in influencing the predicted water vapor distribution in NWP models.

2. EVALUATION OF THE ECMWF HUMIDITY ANALYSES

In the earlier study, we found that the FGGE IIIb global meteorological analysis

Table 1. June 1979 Experimental Mean Error (EME) of Specific Humidity in Percent Bias (QP in Percent) for Those Region/Level Categories Where Systematic Error Occurred in at Least Seven of Ten 12-Hour Forecast Intervals.

Pressure (hPa)	PL GSM Experiment				
	PL-91	SFEVP	QNORM	EMANU	PLKUO
			90S-30S		
850					
700			21		
500	37	-3	33	28	33
400	49	-2	45	40	44
300	82	14	68	75	82
			30S-30N		
950	-9	3	-7		7
900	-12	2	-11		
850	-4	0	-5		
800	-8	-3	-10	-11	-16
700					
600					
500					
400					16
300	15	43			36
			30N-90N		
950				13	15
900				14	17
850				11	15
800					10
700				4	6
600					8
500	16	22	14	13	16
400	16	15	15	11	15
300	40	28	39	24	28

Table 2. Same as in Table 1 But for Relative Humidity Experimental Mean Error (Percent).

Pressure (hPa)	PL GSM Experiment				
	PL-91	SFEVP	QNORM	EMANU	PLKUO
			90S-30S		
850					
700	9	-3	9		
500	11	-6	12		11
400	14	-2	12	12	12
300	23	10	17	21	23
			30S-30N		
950	-17	-4	-16	-7	3
900	-18	-6	-18	-11	
850	-12	-7	-12	-11	
800	-12	-8	-13	-15	-10
700			-4	-6	
600	-7	-6	-8	-8	
500	-7	1	-8	-5	
400			-4		10
300	10	19	8	11	17
			30N-90N		
950		-7	-6		
900	-5	-4	-6		
850	-6	-4	-7	-3	
800	-6	-1	-7	-4	
700	-5	-3	-6	-4	
600					
500	4	9	4	4	6
400	6	8	6	6	7
300	16	12	15	13	12

data sets for January and June 1979 were characterized by a significant systematic positive specific humidity bias below 300 mb. Although we imposed a correction on the analyses to remove most of this bias, we do not know how this alteration may have affected model feedbacks that could influence the forecasts. In particular, the process used to produce these analyses (described by Uppala[12]) involved an initialization procedure that influences the humidity fields. Sheu and Curry[13] observed that in the January 1979 FGGE IIIb analyses in a region over the northern Atlantic Ocean, the initialized analyses were 20% more moist below 850 mb and 20% drier above 500 mb than the uninitialized analyses. Because of the unknown influence of the initialization procedure, we are not certain if the initialized humidity analyses were formed in such a way as to especially optimize the performance of the European Centre for Medium Range Weather Forecasts (ECMWF) forecasts.

For this reason, I acquired uninitialized global meteorological analyses from the ECMWF/WCRP Level III-A Global Atmospheric Data Archive for this follow-on study. In particular, I obtained the ECMWF/TOGA Advanced Operational Analysis Upper Air Data Set for January, February, July, and August 1991. These data include geopotential, temperature, vertical velocity, zonal and meridional wind components, and relative humidity on 14 constant pressure levels. The data sets were in the form of spherical harmonics at a triangular 106 wave (T106) spectral resolution, for 0000 and 1200 UTC times. The purpose for acquiring these data sets was to determine the sensitivity of the various altered forms of the PL GSM to a more modern set of analyses that more closely conform to the observed water vapor distribution (at least in RAOB data-rich areas).

I first sought to establish the conformity of the ECMWF moisture analyses to RAOB humidity measurements by verifying the analyses against RAOB data. I limited my investigation to the single month of July 1991. The U.S. Air Force Environmental Technical Applications Center (USAFETAC) provided upper air soundings (RAOBS) for this period, which were used for all verifications described in the balance of this report. The twice-daily ECMWF analyses of temperature (T) and relative humidity (RH) rendered on a 1.875° latitude-longitude grid were used to compute specific humidity (q) at all grid points on the 14 reported pressure levels.

I then computed the q monthly mean error and the standard deviation of the individual analysis mean error about the monthly mean error for the verification regions. This was done by horizontally interpolating the gridpoint q values to RAOB locations. Profiles of these statistics, as well as the profiles for the June 1979 FGGE IIIb analyses (which we reported in our previous study[5]), are shown in Figures 1 and 2. The July 1991 ECMWF operational analyses have a systematic positive moisture bias at RAOB locations that is much smaller than the June 1979 FGGE IIIb bias. Because these are uninitialized analyses, this simply means that the analysis process properly assimilated the RAOB humidity observations.

3. PL GSM FORECAST EXPERIMENTS INITIALIZED BY ECMWF ANALYSES

Having confirmed the veracity of the ECMWF humidity analyses (at least at humidity observation locations), I now sought to assess the impact of these analyses on forecasts from the various experimental forms of the PL GSM. As in the previous study, the PL-91 versions and three alternative formulations (SFEVP, EMANU, and QNORM) were integrated using the same 18-layer, rhomboidal 30-wave configuration. The forecasts were run out to five days, and forecast spectral coefficients were saved every 12 hours of forecast time. In this case, the forecasts were initialized from only four initial times: 0000 UTC 1 and 16 July 1991, and 1200 UTC 8 and 23 July 1991.

In the verification process, forecast spectral coefficients were postprocessed to a 1° latitude-longitude grid on eight constant pressure levels for humidity. I then bilinearly interpolated gridded forecasts to each RAOB location and computed $q^f - q^o$. Finally, Eq. (2) was used to compute $IME_{j,n}$ for each individual forecast and Eq. (1) to compute ME_n for the ensemble of forecasts. From these, Eq. (3) is used to compute SDE_n.

Because the sample size of the forecast ensemble is limited to four, the determination of systematic error according to the aforementioned criteria ($|ME_n| \geq SDE_n$) is less strongly statistically supported in this case (as contrasted with the sample of 12 in the previous study). For this reason, I computed and plotted QP separately for each of the four forecasts. This allows a subjective determination of

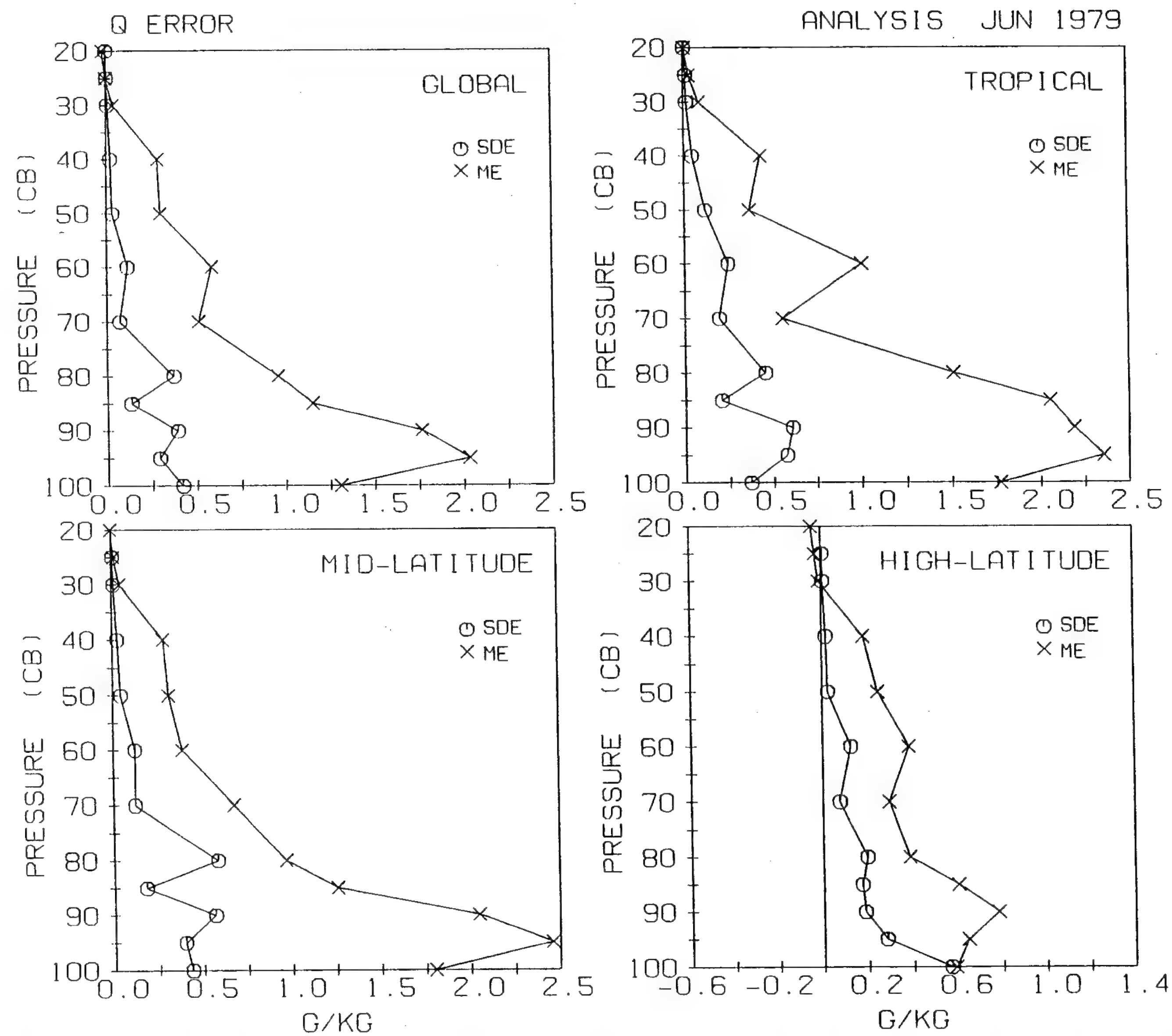

Figure 1. Ensemble Mean Error (ME) of Specific Humidity, and Standard Deviation of the Individual Analysis Mean Error About the ME (SDE) of Specific Humidity, for June 1979 FGGE IIIb Initialized Analyses.

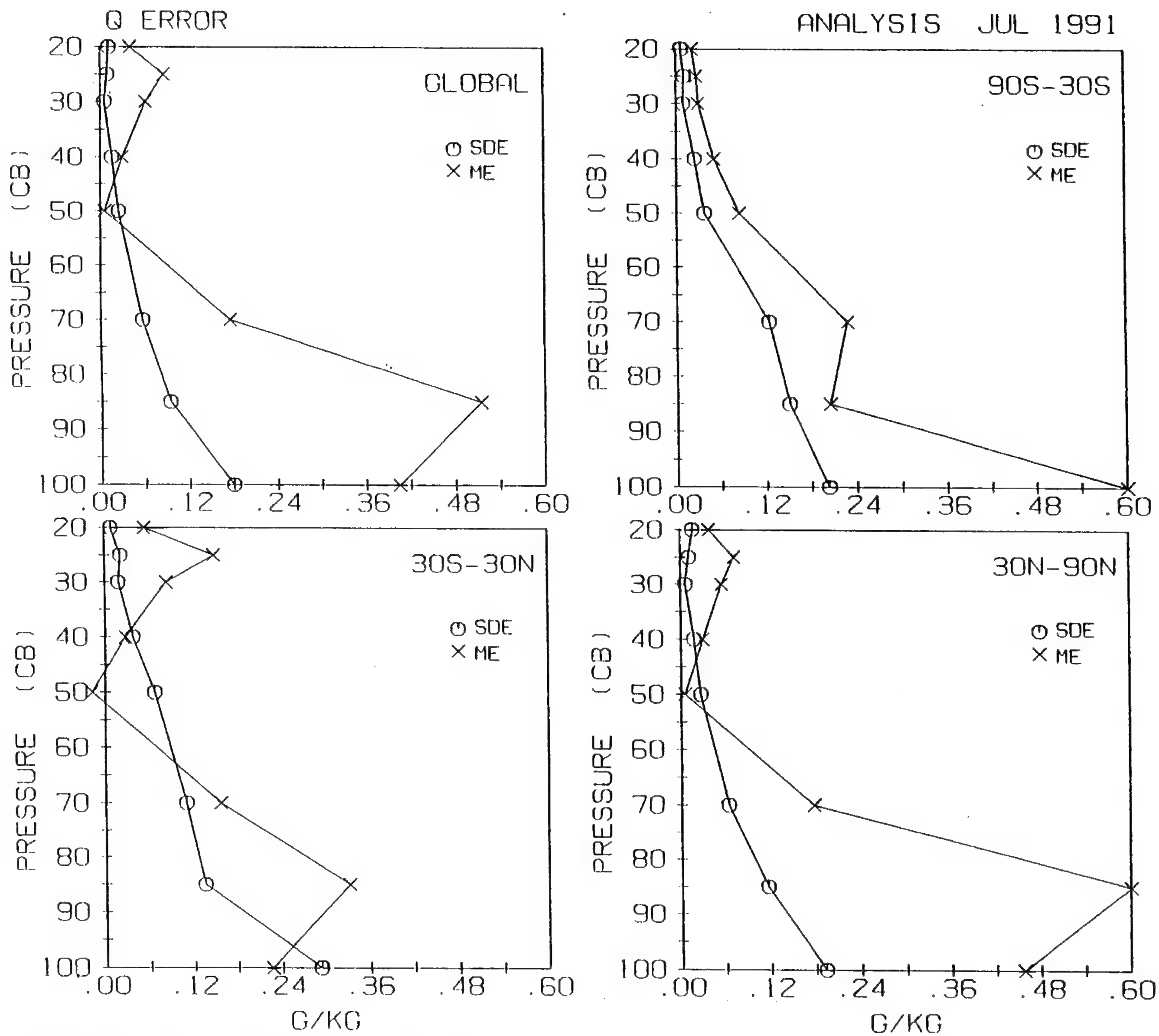

Figure 2. Ensemble Mean Error (ME) of Specific Humidity, and Standard Deviation of the Individual Analysis Mean Error About the ME (SDE) of Specific Humidity, for July 1991 ECMWF Uninitialized Analyses.

the degree of consistency of the bias among the four independent forecasts. Figures 3 and 4 display these plots for the 12- and 48-hour forecast durations (n = 1,4). Though QPs were plotted for 1000, 250, and 200 hPa, they were not included in systematic error assessment. The 1000 mb level had too few RAOB reports to generate reliable statistics, and the RAOB measurements above 300 mb are not considered to be as reliable as those lower in the troposphere. Finally, I computed the experimental mean error (EME) in QP over all RAOBs in the ensemble using Eq. (4). These values of EME are included in Table 3 for those regions/pressure level categories in each experiment where at least seven of the ten 12-hour forecast times have systematic error ($|ME| \geq SDE$) *and* the individual forecast QP is of the same sign in all four forecasts (as seen in Figure 2).

In comparing Table 3 (July 1991 ECMWF initial conditions) with Table 1 (June corrected FGGE IIIb initial conditions), we see a few general similarities and a number of specific differences. Three of the model versions produced same patterns of systematic moistening errors from both initial conditions: PL-91, QNORM, and EMANU overly moisten the middle and upper levels of both hemispheres' extratropics; EMANU creates a moist bias in the summer hemisphere extratropical lower troposphere while PL-91 and SFEVP create excessive water vapor concentrations at the 300 hPa level in the tropics. There are four differences in the responses of the model versions to the two initial conditions:

(1) All models show considerably stronger middle and upper tropospheric erroneous moistening in both hemispheres' extratropics in July 1991,

(2) tropical low level (pressures ≥ 800 hPa) systematic drying error in PL-91, QNORM, and EMANU, in June 1979 is not present in July 1991,

(3) tropical systematic moistening error not present in June 1979 occurs in July 1991 forecasts at 700 hPa in all four common model versions, which also exists at 400 hPa in PL-91, SFEVP, and EMANU, and at 300 hPa in QNORM and EMANU, and,

(4) a modest 850 and 700 hPa summer extratropical moistening error in July 1991 PL-91, SFEVP, and QNORM forecasts that did not appear in the July 1979 forecasts. Overall, the July 1991 forecasts have larger values of experimental mean error of specific humidity than the June 1979 forecasts in almost every common level and

region of systematic error for all four common model versions.

In comparing the impact of the alternative forecast methods, we find that the differences between the two months are greatest in the SFEVP forecasts. In June 1979, the middle and upper levels of the winter hemisphere extratropics reveal a significant reduction of the PL-91 systematic moistening bias by SFEVP. SFEVP also reduced the PL-91 tropical low level drying. By contrast, the July 1991 SFEVP made no appreciable reduction to the PL-91 systematic error in any region or level. QNORM does reduce systematic error more in July 1991 than in June 1979 in 30N-90N, but not in 30S-30N or 90S-30S (the latter shows considerable June 1979 bias reduction). Extratropical systematic moistening error reduction at 500-300 hPa by EMANU is similar in both months. July 1991 humidity forecasts are more moist with respect to RAOBS than the June 1979 forecasts in almost all region/level/model version categories where systematic error is identified.

The following considerations are possible reasons for differences between bias magnitudes for the two months. First of all, both the months and the years are different. Although June and July should have similar water vapor climate characteristics, the interannual variability is unknown. What is likely to be a more important difference, however, is the specified water vapor distribution differences between the corrected FGGE IIIb analyses of June 1979 and the ECMWF analyses of July 1991. We compare the zonal, monthly average percent specific humidity differences between original FGGE IIIb and the aforementioned RAOB-based climatology[6] with the ECMWF--climatology differences in Figure 5. Notice that below 50 kPa, the original FGGE IIIb analyses of June 1979 are in the monthly, zonal mean, much more moist than the ECMWF July 1991 analyses. In the corrected FGGE IIIb analyses (used to initialize the June 1979 "Cor" forecasts), the zonal, monthly percent specific humidity differences from the climatology are zero at all levels and latitudes. Thus, in water vapor concentration, original FGGE IIIb (June 1979) > ECMWF (July 1991) > corrected FGGE IIIb (June 1979). The more moist initial conditions of the July 1991 forecasts may help explain their greater systematic moist bias magnitudes. Indeed, our prior study[5] revealed somewhat greater moist biases when PL-91 was initialized from original (moister) FGGE IIIb analyses than

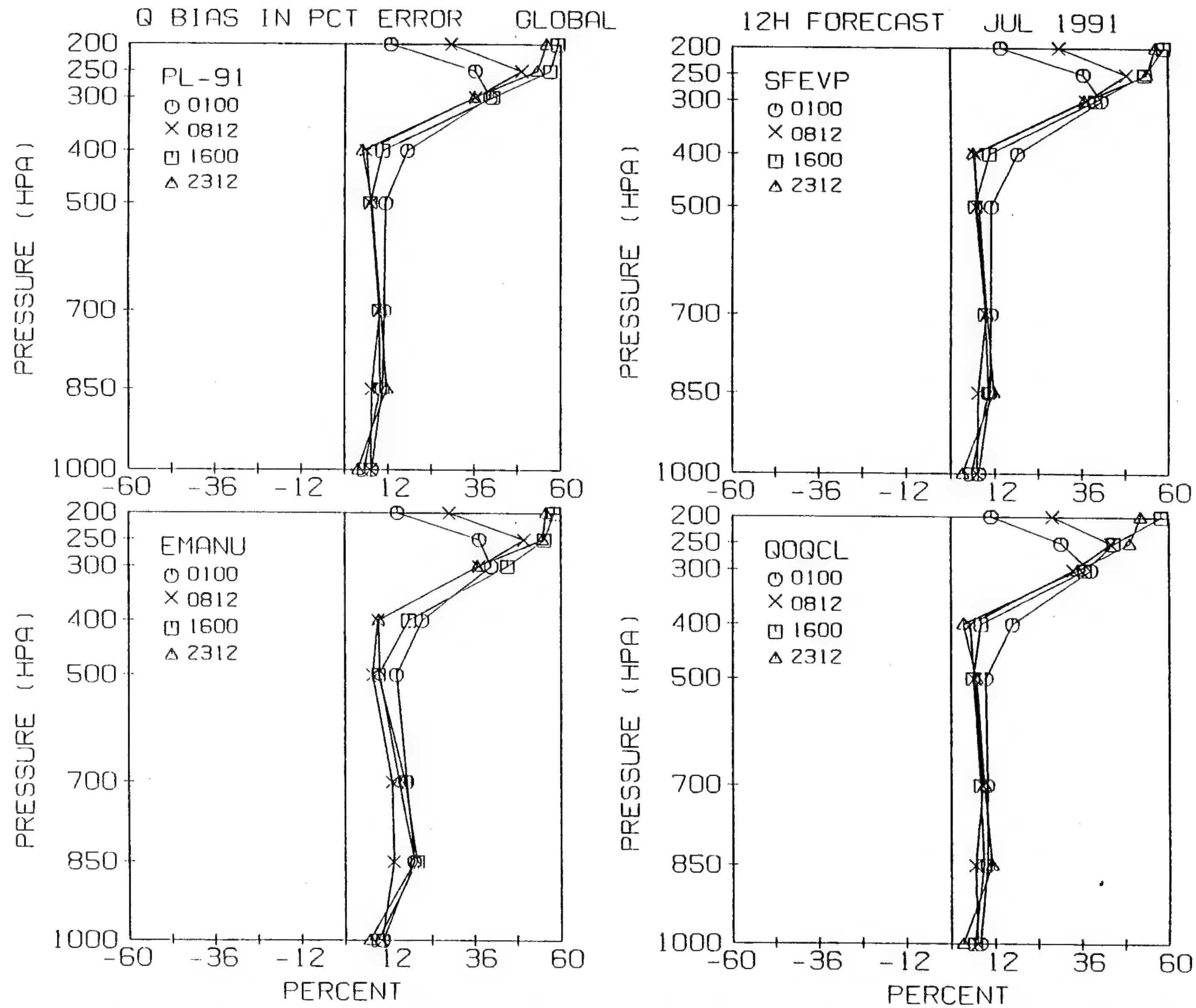

Figure 3. Individual Forecast Mean Errors of Specific Humidity (in Percent Error) for Four July 1991 Forecasts From Each of Four Versions of the PL GSM for 12-Hour Forecasts.

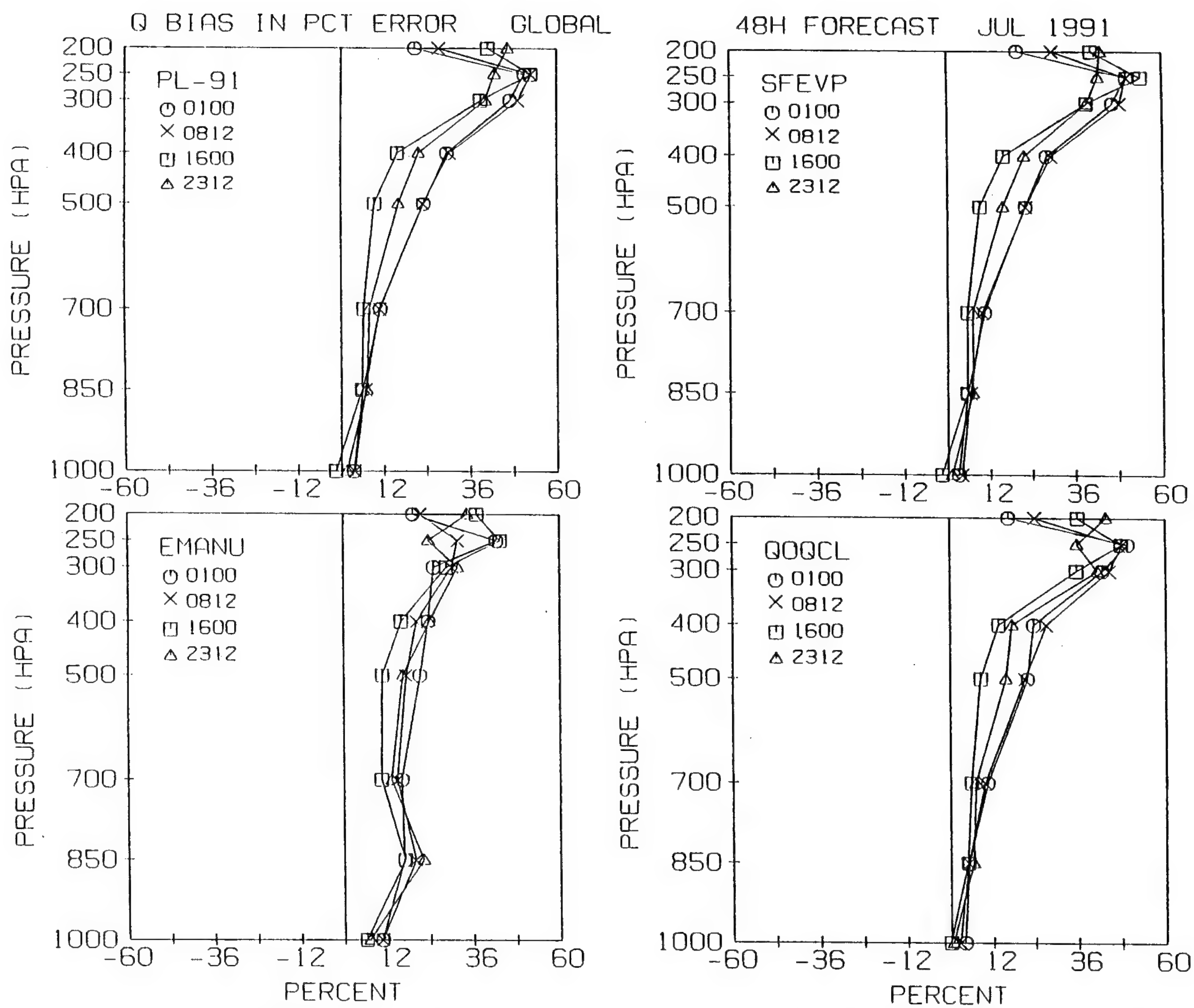

Figure 4. Individual Forecast Mean Errors of Specific Humidity (in Percent Error) for Four July 1991 Forecasts From Each of Four Versions of the PL GSM for 48-Hour Forecasts.

Table 3. July 1991 Experimental Mean Error (EME) of Specific Humidity in Percent Bias (QP in Percent) for Those Region/Level Categories Where Systematic Error Occurred in at Least Seven of Ten 12-hour Forecast Intervals.

Pressure (hPa)	PL GSM Experiments			
	PL-91	SFEVP	QNORM	EMANU
		90S-30S		
850				
700	47	46	47	37
500	73	72	73	57
400	87	86	87	73
300	151	150	145	133
		30S-30N		
850				10
700	10	10	8	9
400				
400	15	15		23
300	32	33	27	30
		30N-90N		
850	8	8	7	20
700	8	7	8	14
500	23	22	21	21
400	29	28	25	23
300	54	52	50	35

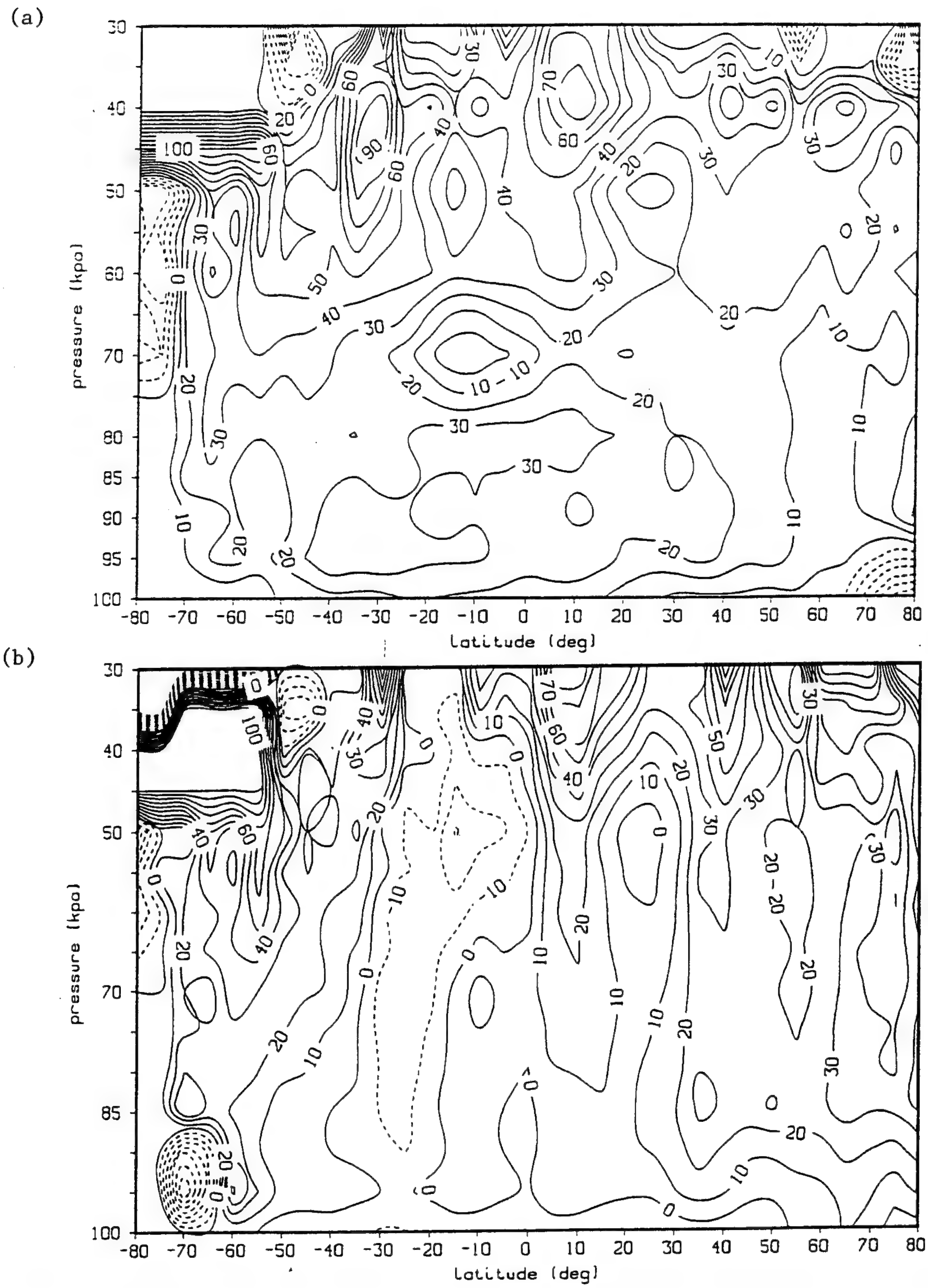

Figure 5. Zonal Cross Sections of Difference (in Percent) Between (a) June 1979 FGGE IIIb Analyses and a RAOB Climatology, and (b) July 1991 ECMWF Analyses and the RAOB Climatology.

corrected analyses. However, the differences weren't as great as the differences seen here between June 1979 and July 1991 forecasts. In this case, not only was the specific humidity specification different between the two analyses (as in the original and corrected FGGE IIIb), but also the specification of mass and motion fields. Due to the complicated feedbacks in the model, there is no way to quantitatively attribute the significantly greater moist biases in the July 1991 to different mass and motion initial conditions. All we can say is that the combined mass, motion, and moisture specification differences between the June 1979 and July 1991 forecast initial conditions contributed to the larger systematic moist biases seen in the July 1991 forecasts. We could only determine how much effect just the humidity differences made if we "corrected" the ECMWF humidity analyses and used them in another set of July 1991 experiments. Finally, there may be unknown differences in the verification data set characteristics between the two months. Various countries have made changes to instruments and reporting practices between 1979 and 1991 which may contribute to the moist bias differences.

To account for differences in alternate model version impacts on PL-91 humidity bias between the two months, we examine atmospheric water sources and sinks for several experiments. Figures 6 and 7 show plots of the globally averaged evaporation (source) and precipitation (sink) for the averages over four June 1979 and July 1991 forecasts. In June 1979, we used forecasts initialized 7½ days apart beginning on 31 May at 1200 UTC, to match the number and time spacing of the July 1991 forecasts. I plotted evaporation and precipitation rates for PL-91 and SFEVP for both months because the differences in impact on the systematic moist bias between the two months was greatest for SFEVP. I looked at the evaporation and precipitation rates for both model versions in both months, to try to understand why SFEVP would have such a substantial impact on the PL-91 moist bias in June 1979 and virtually no impact in July 1991. For quantitative reference, five day average rates and the monthly climatological rates are included in Table 4.

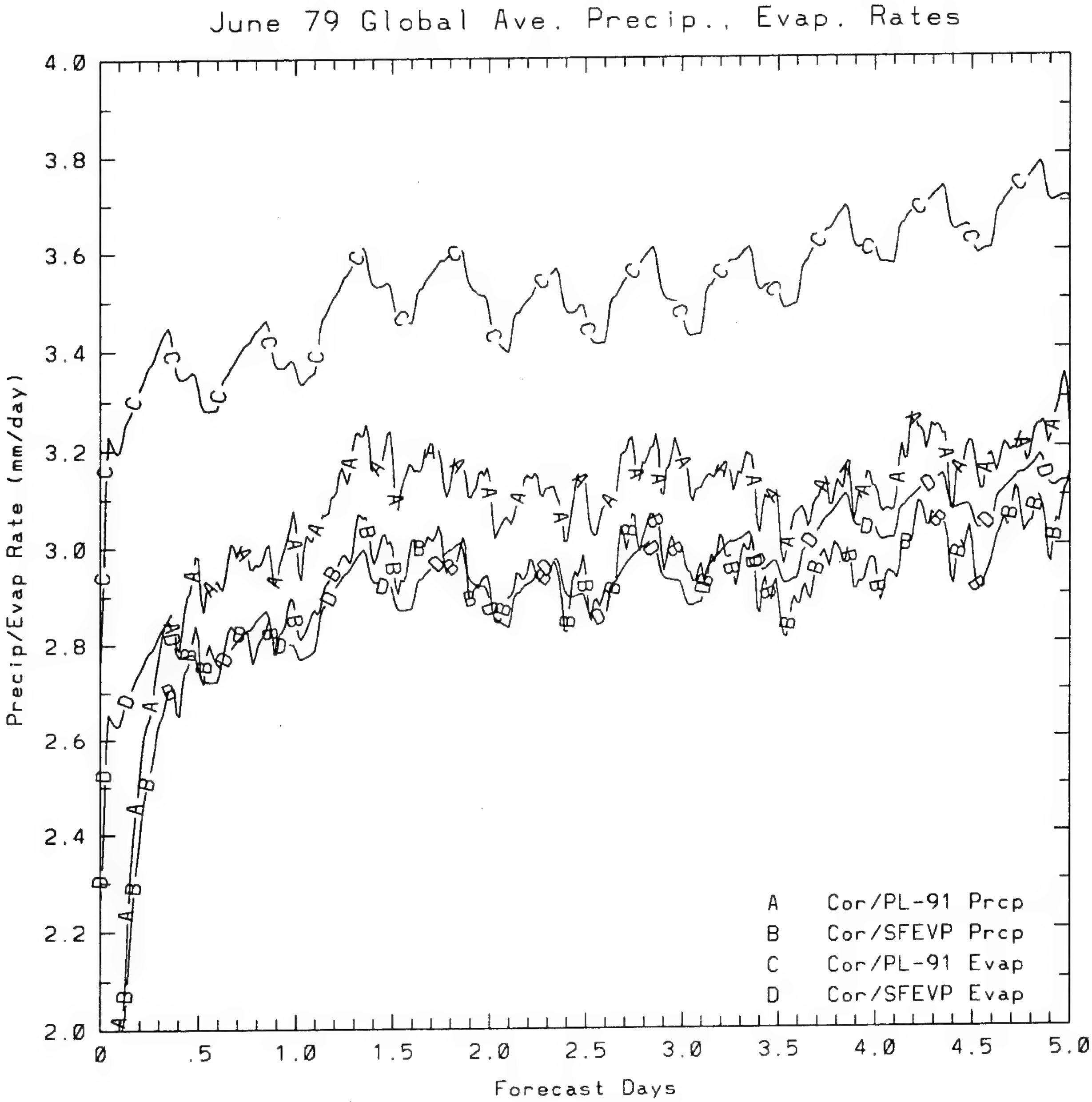

Figure 6. Globally Averaged Precipitation and Evaporation Rates (mm/day) for an Ensemble of Four PL-91 and SFEVP Forecasts from June 1979 FGGE IIIb Analyses.

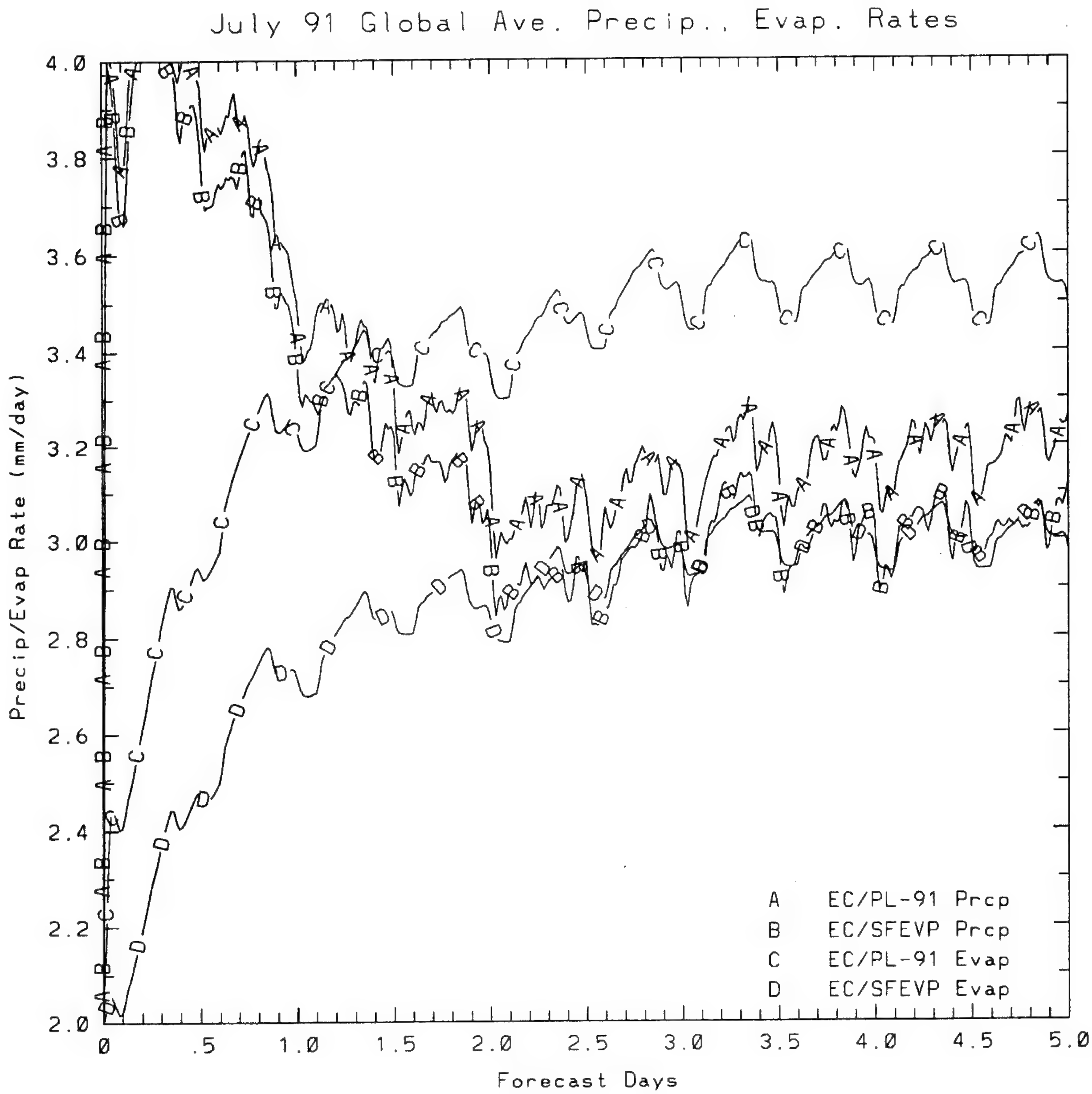

Figure 7. Globally Averaged Precipitation and Evaporation Rates (mm/day) for an Ensemble of Four PL-91 and SFEVP Forecasts from July 1991 ECMWF Analyses.

Table 4. Global Average Monthly Mean Precipitation and Evaporation Rates Averaged Over Five Days of Forecast Time For Several PL GSM Experiments, and Climatological Monthly Mean Precipitation Rates (all rates in mm/day).

Experiment	Precipitation Rate	Evaporation Rate	Precipitation/Evaporation
Cor/PL-91	3.06	3.52	0.87
Cor/SFEVP	2.89	2.04	0.98
June Clim.	2.97		
EC/PL-91	3.35	3.37	0.99
EC/SFEVP	3.20	2.85	1.12
July Clim.	2.84		

In Figures 6 and 7, we see that for PL-91, evaporation rates exceed precipitation rates throughout the five-day forecast in June 1979 and over the last four days in July 1991. Since PL-91 precipitation rates exceed climatological rates in both months, we can be sure the evaporation rates are excessive. In June 1979, the evaporation rates start out strong and grow stronger in the five days. In July 1991, evaporation rates take a little over one day to reach initial June 1979 levels. Meanwhile, the June 1979 precipitation rates "spin up" in the first day, while the July 1991 precipitation rates decrease from their initial excessive values. The differences in the first day between June 1979 and July 1991 PL-91 rates are attributable to the more humid initial conditions of the July 1991 forecasts (see Figure 5). The drier June 1979 initial conditions immediately lead to higher evaporation rates as the model seeks to "re-moisten" the atmosphere to the model's equilibrium water vapor state. Even so, the June 1979 precipitation rates lag behind and equilibrium between source and sink is not achieved by five days of forecast time. This same imbalance takes place after the first day In the July 1991 forecasts (in which the precipitation mechanisms "rain out" excessive water vapor), suggesting that PL-91 evaporation rates are excessive regardless of the initial conditions. That is, even when the initial conditions are somewhat too moist with respect to RAOB data [see Figures 2 and 5], the PL-91 evaporation rates are excessive.

The SFEVP model version has the effect of bringing precipitation and evaporation rates into balance by ½ day in June 1979 and by 2 days in July 1991. This is mostly

due to the reducing effect that SFEVP has on the evaporation rates from their PL-91 levels (again, see Figures 6 and 7, and Table 4). The near balance between precipitation and evaporation rates for Cor/SFEVP (ratio nearly 1 in Table 4) as compared to the precipitation/evaporation imbalance for Cor/PL-01 (ratio much less than 1.0 in Table 4) seems to be associated with the reduction of systematic moist bias by SFEVP in the June 1979 forecasts (see Table 1). However, for the July 1991 SFEVP forecasts, we don't see the relationship between an increase in the precipitation/evaporation ratio (now greater than 1, indicating net atmospheric drying) and a reduction in systematic moist bias taking place.

Why do we see reduction of systematic moistening from PL-91 levels in SFEVP for June 1979 forecasts, but not for July 1991 forecasts, even though we see increases of precipitation/evaporation ratios of SFEVP forecasts in both months? Though a definitive answer cannot be ascertained from these results, one explanation lies in the fact that the June 1979 forecasts began from an altered humidity specification. It may be that the improvement in humidity bias in SFEVP in this case was due to the reduced evaporation rates, preventing the model from remoistening the atmosphere as we saw in PL-91. In the July 1991 forecasts, the model atmosphere was already near the water vapor equilibrium state, so any reduction in the evaporation rate did not negate the already overly moist atmosphere that the model integration began with. Thus, in the July 1991 forecasts, I speculate that initial conditions played more of a role in the final moisture state of the atmosphere than model formulation did. Substantiation for this speculation might be attained by running PL-91 and SFEVP experiments on the original FGGE IIIb initial conditions to see if they show a lesser effect for SFEVP. I did not do this in this project. One clear benefit of the corrected humidity analyses as initial conditions is the balance attained by precipitation and evaporation at about the climatological level for precipitation. This, added to the improvement in temperature forecasts (as seen in the earlier study[5]) and reduced humidity bias (with SFEVP), suggests a significant benefit to properly specified humidity in initial conditions of the forecast.

4. HUMIDITY FORECAST BIAS IN OPERATIONAL MESOSCALE NWP MODELS

The Air Force is currently in the process of selecting a mesoscale NWP model to field in the Combat Weather System (CWS). The CWS is envisioned to be an ensemble of meteorological analysis, forecast, and display software to be operated on a high-powered workstation. An essential analysis and forecast parameter to be generated by the CWS is atmospheric water vapor distribution. The object of this research effort was to provide information to the Air Force decision-makers on the systematic humidity forecast errors typical of operational mesoscale models so that they can be prepared for the general capabilities of current models in predicting humidity. It is unlikely that the magnitude and distribution of systematic humidity error will be a decisive factor in the selection of the CWS model. However, such knowledge can be useful in planning possible modifications for either the forecasts or the model formulation itself, of the model selected.

To provide such information, I requested "off-the-shelf" absolute humidity bias (mean error) statistics for mesoscale models executed regularly at several operational weather forecast centers. I asked that the statistics be based on the verification of the model's forecasts with radiosonde data over a generally data-rich region, for both a winter and summer month. As in the current study, I did not expect to include verification statistics on pressure levels corresponding to pressures less than 300 hPa.

In our earlier global model study, we were able to determine whether the mean error, computed over an ensemble of forecasts, was indeed systematic. We did not expect that the providers of the information we requested for the mesoscale models would provide the individual forecast verifications for an entire month. Rather, we asked only for the monthly ensemble mean errors. Therefore, it is not possible from the statistics obtained to determine with certainty if the humidity prediction has systematic error. In fact, over a limited region, it is probable that the regional mean error will vary more greatly from forecast-to-forecast than in the case of evaluating the mean error over ¼ (extratropics) to ½ (tropics) of the globe as we did in our study. In this discussion, I can only present monthly mean error without regard to the

forecast-to-forecast consistency of the error. The following sub-sections discuss the error statistics received in response to my request.

4.1 Humidity Verification Statistics from Air Force Global Weather Central (AFGWC) Relocatable Window Model (RWM)

AFGWC executes its RWM regional NWP model twice-daily over several fixed and contingency regions (referred to as "windows"). The model runs on a 61X61 grid at 50 nm grid spacing (hence, a window is 3000 nm on a side), with 16 model layers in the vertical. The numerical and physical parameterization formulation of the RWM is based on the NMC's Quasi-Lagrangian Nested Grid Model.[14]

Table 5 lists the monthly mean error of temperature (T), relative humidity (RH), and specific humidity (q) forecasts from the RWM for January and July 1993 over the "United States" window. This window is centered on the continental U.S., and includes southern Canada, northern Mexico, and portions of the Pacific and Atlantic Oceans near the U.S. coast. Radiosonde (RAOB) data were used for the verification, after the forecast fields were interpolated horizontally to each RAOB location, and vertically to mandatory pressure levels.

According to Table 5, the RWM produces a positive q bias (water vapor concentrations too high) on the average at all levels in both months except at 850 and 1000 hPa in July. At these two levels, water vapor concentrations are too low, compared to observations. In January, at 700-300 hPa, the moist bias increases slowly with forecast duration. In July, at 850 and 1000 mb, the dry bias grows steadily with increasing forecast length. The monotonic nature of these error changes with forecast length suggest that these errors may indeed be systematic; however, this cannot be stated with certainty. There is no appreciable growth of the q bias with forecast length at lower levels in January and at upper levels in July.

Table 5. Monthly Mean Error of Relocatable Window Model Temperature, Relative Humidity, and Specific Humidity Forecasts over the United States Window When Verified Against RAOBs

January 1993

Pressure(mb)	Temperature (K)			Rel. Humidity			Spec. Humidity (g/kg)		
	12H	24H	36H	12H	24H	36H	12H	24H	36H
300	0.05	0.31	1.03	0.09	0.16	0.16	0.05	0.07	0.08
400	0.10	0.46	1.12	0.09	0.16	0.18	0.07	0.13	0.16
500	0.18	0.52	1.07	0.06	0.11	0.18	0.07	0.13	0.16
700	0.32	0.81	1.31	0.02	0.02	0.02	0.18	0.29	0.35
850	0.38	0.96	1.47	-0.03	-0.06	-0.07	0.02	0.02	0.03
1000	0.75	0.63	0.52	-0.01	-0.01	0.01	0.29	0.27	0.29

July 1993

Pressure(mb)	Temperature (K)			Rel. Humidity			Spec. Humidity (g/kg)		
	12H	24H	36H	12H	24H	36H	12H	24H	36H
300	-0.36	-0.38	-0.62	0.06	0.06	0.11	0.04	0.03	0.06
400	-0.61	-0.58	-0.79	0.06	0.07	0.08	0.09	0.08	0.10
500	-0.52	-0.42	-0.54	0.04	0.04	0.05	0.07	0.05	0.10
700	0.04	0.23	0.39	0.00	-0.01	-0.01	0.13	0.13	0.09
850	-0.27	-0.21	-0.03	-0.04	-0.06	-0.07	-0.58	-0.82	-1.01
1000	-0.21	-0.37	-0.38	-0.07	-0.08	-0.09	-1.16	-1.50	-1.70

The relationships between the T and q biases in Table 5 illustrate a danger in drawing conclusions about NWP water vapor forecast error from a measure of saturation, like RH. When both T and q biases are of the same sign, the effect of erroneous temperature and water vapor concentrations are offsetting in the calculation of RH. The clearest examples of this effect in Table 4 are the erroneously cold temperatures at 850 and 1000 hPa in July masking the significantly dry specific humidities, resulting in a relatively modest underprediction of relative humidity. If RH is the only moisture variable verified in model forecasts, the likely low level systematic drying of the model in summertime forecasts may remain undiscovered and uncorrected. In the opposite sense algebraically, in January at 700-300 hPa, an appreciable warming in the model obscures a steadily growing moistening, leading

to RH that appears to be leveling off in its growth rate by 36 hours of forecast time. To avoid drawing incorrect conclusions about the trends in water vapor forecasts by a model, absolute humidity errors should be examined.

4.2 Humidity Verification Statistics from the NMC Eta and LFM

The National Meteorological Center has replaced its long-running Limited-area Fine-mesh Model (LFM) with the recently-developed Eta model. The latter model and the ways in which it differs from LFM are described by Black et al.[15] This report also includes verification statistics for the "Early", or 80 km resolution version of the Eta model, along with statistics for LFM forecasts. Eta model and LFM forecasts were verified against North American RAOBs during the period 26 February-1 April 1993. Figure 8 is a graphical presentation of specific humidity bias of forecasts of Eta and LFM during this verification period, taken from the Black et al.[15] report.

Both models' forecasts show a substantial and consistent bias at 1000 mb (same as hPa). Water vapor concentrations are too high in Eta forecasts and too low in LFM forecasts. At 850 and 700 mb, the moist bias in Eta forecasts is considerably less than in LFM forecasts. A trend we see in 850 and 400 mb Eta forecasts is a decrease in the magnitude of the q bias with forecast duration. At 500 mb, a decrease in dry bias gives way to a moist bias by 48 hours, suggesting a trend toward moistening at this level. At 400 and 300 mb, Eta and LFM models exhibit a dry bias, unlike the consistent moist bias we saw in RWM forecasts (Table 5). These model differences accent the fact that the nature (sign and magnitude) of water vapor forecast errors can vary greatly in its vertical distribution from model to model. Hence, the cure for the systematic error in one model may not fix the problem in another model.

The NMC Nested Grid Model[16,17] (NGM) has been in operation for about 10 years, producing twice-daily forecasts over North America. In response to my request, NMC personnel sent graphs of 850 and 500 hPa mean error statistics of T and RH forecasts for January and July 1993. In this case, I do not have absolute humidity forecast verification statistics. Thus, I cannot draw firm conclusions about the nature of water vapor forecasts from the NGM. Also, having only data from two pressure

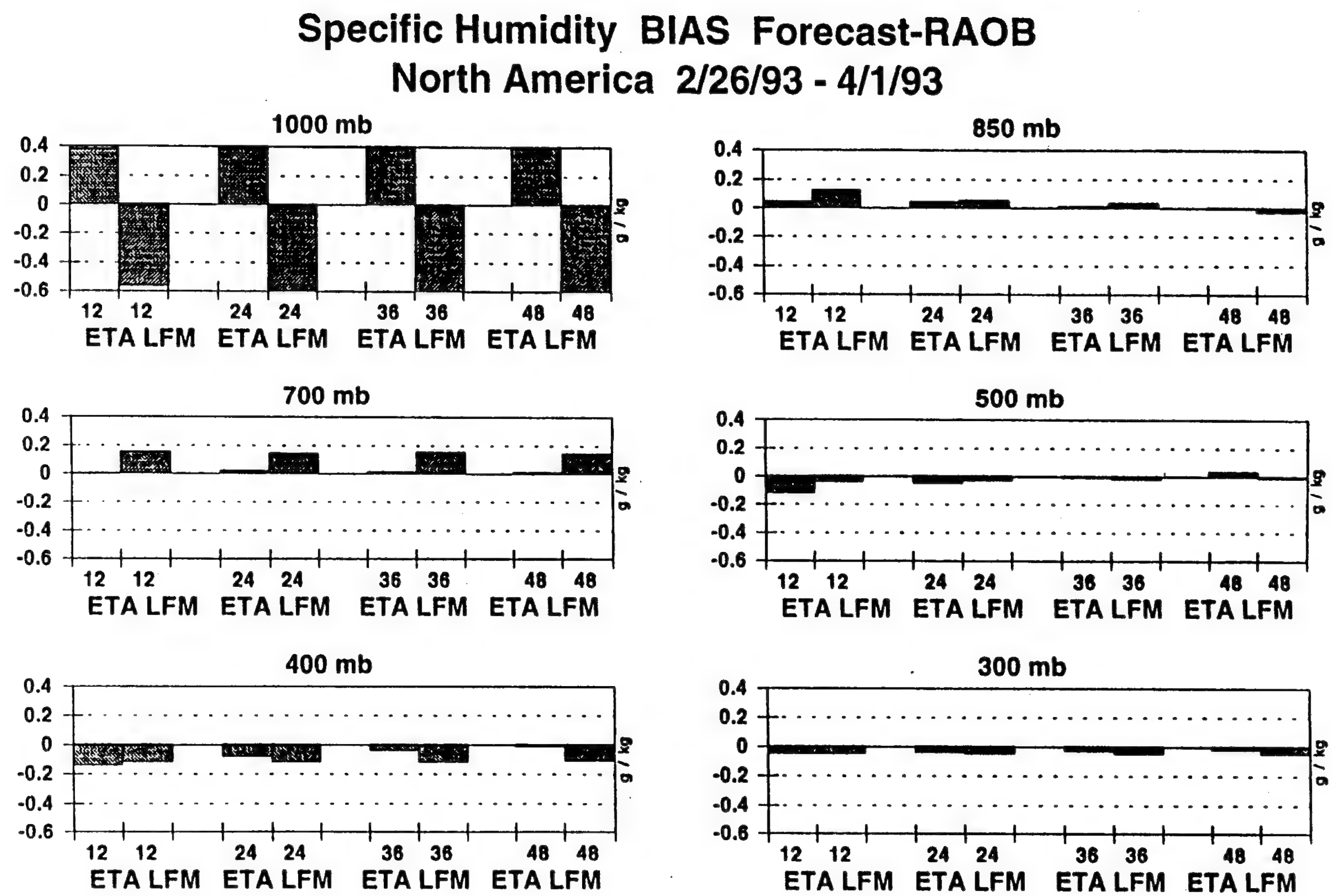

Figure 8. Specific Humidity Bias (q/kg) of a Series of NMC Eta Model and LFM Forecasts Computed for the Period 26 February-2 April 1993 for 12-48-Hour Forecasts.

levels precludes a discussion of the vertical distribution of the humidity forecast error. Table 6 lists the monthly mean error statistics extracted from the graphs supplied by NMC.

Table 6. Monthly Mean Error of Nested Grid Model Temperature and Relative Humidity Forecasts Over North America When Verified Against RAOBs.

	January 1993				July 1993			
Forecast Duration	Temperature(K)		Rel. Humidity		Temperature(K)		Rel. Humidity	
	850 mb	500 mb	850 mb	500 mb	850 mb	500 mb	850 mb	500 mb
0H	0.1	0.1	0.02	0.01	0.7	0.8	0.00	0.01
12H	0.2	-0.1	0.02	0.05	0.1	0.4	0.00	0.04
24H	0.2	-0.1	0.02	0.08	-0.5	0.2	0.02	0.05
36H	0.2	-0.1	0.02	0.10	-1.1	-0.2	0.02	0.07
48H	0.1	-0.1	0.01	0.11	-1.5	-0.8	0.03	0.08

In the NGM statistics, we see that the modest RH biases at 500 mb seem to grow with forecast duration. Given the small magnitude and consistency of the corresponding temperature errors in the January 500 mb forecasts, it is likely that the positive RH biases signal a growing moist bias in water vapor forecasts. Such a bias in the July 500 mb forecasts is not certain.

5. CONCLUSIONS

This report describes the systematic humidity bias of global meteorological analyses, and of forecasts initialized from them produced by the Phillips Laboratory Global Spectral Model. I found that both the specified initial moisture conditions in the analyses and the forecast model formulation can affect the sign and magnitude of the forecast humidity bias. One clear conclusion from this work is that the impact of initial conditions or model formulation is not transferable between differing analyses or forecast models. One must diagnose the nature of the humidity bias in his/her own analysis and forecast suite, before drawing conclusions as to what must be done to reduce the systematic humidity error in both. This report has defined systematic error in such a way that such error in humidity analysis and forecasts can

be readily discerned. Steps in identifying the causes of such error in the forecast model are described in an earlier paper.[5] However, this report has shown that forecast model remedies, with respect to one set of initial conditions, do not guarantee equally reduced moisture biases when the model is initialized from a different analysis data set.

In comparing the humidity forecast errors from several mesoscale models, I found appreciable differences in the nature of the errors. This suggests that, like global models, mesoscale models are likely to have unique forecast error characteristics. I also found that absolute humidity must be examined to determine the true nature of the systematic humidity error, because temperature error can mask water vapor errors when just relative humidity is examined.

References

1. Heckley, W.A. (1985) *The Performance and Systematic Errors of the ECMWF Tropical Forecasts*, ECMWF Tech. Rep. No. 53, 97 pp.

2. Williamson, D.L. (1990) Semi-Lagrangian Moisture Transport in the NMC Spectral Model, *Tellus*, **42A:** 413-428.

3. Campana, K., Mitchell, K., Yang, S., Hou, Y. and Stowe L. (1991) Validation and Improvement of Cloud Parameterization at the National Meteorological Center, *Preprints, 9th Conference on Numerical Prediction*, 14-19 October 1991, Denver, CO, pp. 131-134.

4. Kuo, Y.-H., Guo, Y.-R, and Westwater E.R. (1993) Assimilation of Precipitable Water Measurements into a Mesoscale Numerical Model, *Mon. Wea. Rev.*, **121:** 1215-1238.

5. Norquist, D.C. and Chang, S.S. (1994) Diagnosis and Correction of Systematic Humidity Error in a Global Numerical Weather Prediction Model, *Mon. Wea. Rev.*, **122:** 2442-2460.

6. Norquist, D.C., Yang, C., Chang, S., and Hahn, D.C. (1992) *Phillips Laboratory Global Spectral Numerical Weather Prediction Model*, PL-TR-92-2225, Phillips Laboratory, Hanscom AFB, MA, 154 pp. ADA267293.

7. Peixoto, J.P., and Oort, A.H. (1983) The Atmospheric Branch of the Hydrological Cycle and Climate, *Variations in the Global Water Budget*, Street-Perrott, A., Beran, M., Ratcliffe, R., Eds., D. Reidel, Dordrecht, pp. 5-65.

8. Jaeger, L. (1983) Monthly and Areal Patterns of Mean Global Precipitation, *Variations in the Global Water Budget*, Street-Perrott, A., Beran, M., Ratcliffe, R., Eds., D. Reidel, Dordrecht, pp. 129-140.

9. Tiedtke, M. (1989) A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models, *Mon. Wea. Rev.*, **117:** 1779-1800.

10. Emanuel, K.A. (1991) A Scheme for Representing Cumulus Convection in Large-Scale Models, *J. Atmos. Sci.*, **21:** 2313-2335.

11. Norquist, D.C. and Yang, C.-H. (1990) *Refinement and Testing of the Moist Convective Parameterization in the GL Global Spectral Model*, GL-TR-90-0285, Geophysics Laboratory, Hanscom AFB, MA 123 pp. ADA 241684.

12. Uppala, S. (1986) The Assimilation of the Final Level IIB Data Set at ECMWF, Part I. Preprints, *Scientific Results of the First GARP Global Experiment*, 14-17 January 1986, Miami, Amer. Meteor. Soc., 24-29.

13. Sheu, R.-S., and Curry, J.A. (1992) Interactions Between North Atlantic Clouds and the Large-Scale Environment, *Mon. Wea. Rev.*, **120:** 261-278.

14. Mathur, M.B. (1983) A Quasi-Lagrangian Regional Model Designed for Operational Weather Prediction, *Mon. Wea. Rev.*, **111:** 2087-2098.

15. Black, T., Deaven, D., and DiMego, G. (1983) *The Step-Mountain Eta Coordinate Model: 80 km 'Early' Version and Objective Verifications*, Technical Procedures Bulletin, Series No. 412, National Service, Office of Meteorology, Program Requirements and Development Division, Silver Spring, MD 20910, 31 pp.

16. Phillips, N. (1979) *The Nested Grid Model*, NOAA Tech Rep., NWS 22, 80 pp.

17. Hoke, J. (1984) Forecast Results for NMC's New Regional Analysis and Forecast System, Preprints, *10th Conference on Weather Forecasting and Analysis*, Clearwater Beach, Amer. Meteorol. Soc., 418-423.